



DOCUMENT de RECHERCHE

« Modélisation SEMIFARMA-HYGARCH
de la persistance du cours du Dow Jones »

Mohamed CHIKHI
Anne PEGUIN-FEISSOLLE
Michel TERRAZA

DR n°2012-06

Unité de Formation et de Recherche d'Economie
Avenue Raymond DUGRAND C.S. 79606
34960 MONTPELLIER Cedex 2

E-mail : lameta@lameta.univ-montp1.fr
web : www.lameta.univ-montp1.fr

MODELISATION SEMIFARMA-HYGARCH DE LA PERSISTANCE DU COURS DU DOW JONES

MOHAMED CHIKHI¹

ANNE PEGUIN-FEISSOLLE²

MICHEL TERRAZA³

Janvier 2012

Résumé

Cet article analyse le comportement cyclique du cours du Dow Jones et notamment ses propriétés de mémoire longue à travers une nouvelle classe de modèles ARFIMA semi-paramétriques avec erreurs GARCH hyperboliques, notée SEMIFARMA-HYGARCH ; cette classe inclut une tendance déterministe non paramétrique, une tendance stochastique, la dépendance à court et à long terme ainsi que le terme d'erreur hétéroscédastique à mémoire longue. Nous étudions les rendements journaliers du Dow Jones de 1896 à 2006. Les résultats prédictifs montrent que les chocs informationnels ont des conséquences durables sur la volatilité et que le modèle SEMIFARMA-HYGARCH possède une supériorité évidente sur d'autres modèles pour des horizons longs et/ou courts. Une des conclusions est que l'hypothèse d'efficience faible des marchés financiers semble violée pour la série des rentabilités de Dow Jones étudiées sur longue période.

Mots-clefs : Modèle SEMIFARMA, modèle HYGARCH, tendance déterministe non paramétrique, méthode du noyau, fenêtre, mémoire longue

Classification JEL : C14, C22, C58, G17

Abstract

This paper analyzes the cyclical behavior of Dow Jones by testing the existence of long memory through a new class of semiparametric ARFIMA models with HYGARCH errors (SEMIFARMA-HYGARCH); this class includes nonparametric deterministic trend, stochastic trend, short-range and long-range dependence and long memory heteroscedastic errors. We study the daily returns of the Dow Jones from 1896 to 2006. The forecasting results show that the informational chocks have permanent effects on volatility and the SEMIFARMA-HYGARCH model has better performance on some other models for long and/or short horizons. We conclude in particular that the weak efficiency hypothesis seems violated for the Dow Jones returns on long period.

Keywords: SEMIFARMA model, HYGARCH model, nonparametric deterministic trend, kernel methodology, bandwidth, long memory

JEL Classification : C14, C22, C58, G17

¹ Université de Ouargla & LAMETA/CNRS, Université Montpellier I, Faculté des Sciences Economiques, Espace Richter, Avenue de la Mer, C.S. 79606, 34960 Montpellier Cedex 2, France. Tel.: +33 (0)4.67.15.84.16, Fax.: 33 (0)4.67.15.84.67, E-mail: chikhi@lameta.univ-montpl.fr

² GREQAM/CNRS, Université d'Aix-Marseille 2, Centre de la Charité, 2 rue de la Charité, 13236 Marseille cedex 02, France, Tel: +33 (0)4.91.14.07.70, Fax: +33.(0)4.91.90.02.27, E-mail: Anne.Peguin@univmed.fr

³ LAMETA/CNRS, Université Montpellier I, Faculté des Sciences Economiques, Espace Richter, Avenue de la Mer, C.S. 79606, 34960 Montpellier Cedex 2, France. Tel.: +33 (0)4.67.15.84.20, Fax.: +33 (0)4.67.15.84.67, E-mail: mterraza@lameta.univ-montpl.fr

1. Introduction

Les cours boursiers présentent des propriétés statistiques particulières dont la prise en compte est fondamentale pour aborder le problème de la modélisation. Ainsi, la présence de longue mémoire dans les séries financières a des implications importantes particulièrement concernant la propriété d'efficiences faibles des marchés. En effet, si une série présente de la longue mémoire, ceci implique des autocorrélations significatives entre des observations qui, bien qu'elles soient éloignées dans le temps, peuvent aider à prévoir les rendements futurs, ce qui viole l'hypothèse d'efficiences du marché (voir, entre autres, Fama (1965, 1970 et 1998), Grossman (1976), Jensen (1978), Lillo et Farmer (2004), Christodoulou-Volos et Siokis (2006), Barkoulas et Baum (1997), Gursakal (2010)). La modélisation des persistances dans les séries financières a donc constitué une grande priorité dans le domaine de la recherche économique. Par ailleurs, les séries temporelles exhibent souvent une tendance déterministe ou stochastique. Ainsi, Beran (1999) a proposé un modèle SEMIFAR (*Semiparametric fractional autoregressive*) qui présente l'intérêt de tenir compte à la fois du comportement de court terme de la série au travers des paramètres autorégressifs, du comportement de long terme par le biais du paramètre d'intégration fractionnaire et de la tendance déterministe non paramétrique. Beran et Feng (2002a, 2002b) ont ensuite introduit la partie moyenne mobile dans le modèle SEMIFAR, noté SEMIFARMA (SEMIFAR *Moving Average*). Le processus SEMIFARMA inclut donc le modèle ARIMA et le processus autorégressif fractionnaire (Granger et Joyeux (1980), Hosking (1981)). Cependant, l'hypothèse de bruit blanc sur les résidus du modèle SEMIFARMA ignore la présence d'hétéroscédasticité conditionnelle ; or, les séries financières sont en général caractérisées par une volatilité variable qui peut être modélisée par des modèles de type ARCH (Engle (1982)) ou GARCH (Bollerslev (1986)). La modélisation SEMIFARMA-GARCH proposée par Feng, Beran et Yu (2007) correspond à une représentation spécifique de la non linéarité qui permet une modélisation simple de l'incertitude. Nous allons étendre ce modèle en utilisant la modélisation HYGARCH (*Hyperbolic GARCH*) (Davidson (2004)) qui fournit une mesure directe de la persistance à travers le paramètre d'intégration fractionnaire.

Notre article s'insère dans la lignée de ces travaux. Il est appliqué à l'analyse de la persistance des chocs informationnels et à la recherche d'une éventuelle mémoire longue dans le cours du Dow Jones. L'étude de la rentabilité journalière de ce cours sur longue période, de 1896 à 2006 (30292 observations), nous permet de montrer que le modèle SEMIFARMA-HYGARCH présente une supériorité prédictive importante sur tous les autres modèles proposés pour des horizons longs. Les prévisions issues de ce modèle sont également supérieures à celles du modèle de marche aléatoire, que ce soit à court terme ou à long terme ; en conséquence, l'hypothèse d'efficiences faibles des marchés financiers semble violée pour la série des rentabilités du Dow Jones étudiés sur longue période.

Dans ce qui suit, la section 2 s'intéresse à la présentation du modèle SEMIFARMA-HYGARCH. La section 3 est consacrée à l'étude empirique de la série journalière du Dow Jones ; nous comparons la qualité prévisionnelle des modèles SEMIFARMA-GARCH, SEMIFARMA-FIGARCH, SEMIFARMA-HYGARCH avec celle d'une marche aléatoire. Nous donnons quelques conclusions dans la section 4.

2. Présentation du modèle SEMIFARMA avec erreur HYGARCH

Certains auteurs ont utilisé le modèle HYGARCH ou le modèle SEMIFAR pour modéliser des séries financières, mais toujours séparément. Ainsi, concernant le modèle HYGARCH, Davidson (2004) étudie la dynamique de la volatilité de trois monnaies asiatiques. Tang et Shieh (2006) analysent la value-at-risk dans le cadre de trois indices boursiers que sont le S&P500, le Nasdaq100 et le Dow Jones. Cardamone et Folkinshteyn (2007) s'intéressent à la sensibilité de taux d'intérêt et de taux de change américains. Härdle et Mungo (2008) étudient la value-at-risk de certains indices boursiers. McMillan et Kambouroudis (2009) comparent les prévisions du modèle HYGARCH et d'autres modèles pour 31 indices boursiers. Kasman, Kasman et Torun (2009) modélisent la présence de longue mémoire dans huit indices boursiers de pays d'Europe Centrale et d'Europe de l'Est. Aloui et Mabrouk (2010) évaluent la value-at-risk pour le pétrole et le gaz. Conrad (2010) dérive les conditions de non négativité de la variance conditionnelle dans un modèle HYGARCH et présente deux applications empiriques, l'une concernant le FTSE journalier, l'autre le DAX30. Wei, Wang et Huang (2010) s'intéressent à la volatilité du marché du pétrole. Enfin, Kwan, Li et Li (2011) étudient certaines propriétés théoriques du modèle HYGARCH et donnent une illustration empirique consacrée au taux de change de la monnaie coréenne. Concernant la modélisation de type SEMIFAR, Beran et Ocker (1999a) l'appliquent à la modélisation de taux de change européens et asiatiques ; Beran et Ocker (1999b) et Feng, Beran et Yu (2007) étudient différentes séries boursières.

La modélisation *SEMIFARMA-HYGARCH* que nous utilisons est la suivante (voir Beran (2002a) pour la partie *SEMIFAR* et Davidson (2004) pour la partie HYGARCH). On dit que $\{Y_t\}$ est un processus semi paramétrique fractionnaire avec erreur GARCH hyperbolique, noté *SEMIFARMA-HYGARCH*, s'il vérifie l'équation suivante :

$$\phi(B)(1-B)^{d_1} \{(1-B)^{d_1} Y_t - g(x_t)\} = \theta(B)\varepsilon_t \quad (1)$$

avec :

$$\varepsilon_t = u_t \sigma_t, \quad \sigma_t > 0, \quad u_t \sim iid(0,1) \quad (2)$$

et

$$\sigma_t^2 = \omega^* + \sum_{i=1}^{\infty} \lambda_i B^i \varepsilon_t^2 = \omega^* + \lambda(B) \varepsilon_t^2 \quad (3)$$

où :

$$\lambda(B) = 1 - [1 - \beta(B)]^{-1} \psi(B) \{1 + \alpha[(1-B)^d - 1]\} \quad (4)$$

et

$$\omega^* = \omega(1 - \beta(B))^{-1} \quad (5)$$

où $(1-B)^d = \sum_{k=0}^{\infty} (-1)^k \frac{\Gamma(d+1)}{\Gamma(k+1)\Gamma(d-k+1)} B^k = 1 - \sum_{k=1}^{\infty} c_k(d) B^k$; de plus, $c_1(d) = d$,

$c_2(d) = \frac{1}{2} d(1-d)$, et $\Gamma(\cdot)$ est la fonction gamma.

et $\phi(B) = 1 - \phi_1 B - \phi_2 B^2 - \dots - \phi_p B^p$, $\theta(B) = 1 - \theta_1 B - \theta_2 B^2 - \dots - \theta_q B^q$ et $\beta(B) = 1 - \beta_1 B - \dots - \beta_r B^r$; $\psi(B) = 1 + \psi_1 B + \psi_2 B^2 + \dots + \psi_s B^s$ et $\lambda(B)$ sont des polynômes retard de degrés p, q, r, s et ∞ respectivement et toutes leurs racines sont en dehors du cercle unité du plan complexe. B est l'opérateur de retard, d_1 est un entier : $d_1 \in \{0,1\}$, $x_t = t$ est la tendance et $g : [0, 1] \rightarrow \mathbb{R}$ est la fonction de lissage. Enfin, le processus $\{Y_t\}$ est stationnaire

et inversible, $-\frac{1}{2} < d_2 < \frac{1}{2}$, et u_t est un processus *iid* qui suit une loi normale. Davidson (2004) montre que le processus HYGARCH généralise le modèle FIGARCH (Baillie, Bollerslev et Mikkelsen (1996)). Si $\alpha = 1$, le modèle HYGARCH correspond à un FIGARCH. Le processus est stationnaire si $0 < \alpha < 1$ et non stationnaire si $\alpha > 1$. La mémoire longue intervient dans l'équation de la moyenne (1) (à travers le paramètre d_2) et dans l'équation de la variance (3) (à travers le paramètre d). Il est important de noter que la définition de la mémoire longue est une définition asymptotique. En d'autres termes, ce qui importe est la façon dont les autocorrélations convergent vers zéro lorsque le retard augmente. En résumé, la distinction, en termes d'autocorrélations, entre un processus à mémoire courte et un processus à mémoire longue résulte de la vitesse de convergence géométrique ou hyperbolique des autocorrélations vers zéro.

Enfin, la fonction $g(x_t)$ représente une tendance déterministe non linéaire. L'estimation par noyau dans le cas de régression non paramétrique a été étudiée notamment par Hall et Hart (1990), Ray et Tsay (1997) et Beran (1999) dans le cas d'erreurs à longue mémoire. Nous estimons g par la méthode du noyau à l'aide du modèle suivant :

$$(1 - B)^{d_1} Y_t = g(x_t) + X_t \quad (6)$$

qui équivaut à $Y_t = g(x_t) + X_t$ si $d_1 = 0$ et à $Y_t - Y_{t-1} = g(x_t) + X_t$ si $d_1 = 1$, où X_t est le processus des erreurs stationnaires à mémoire longue si $d_2 > 0$ et à mémoire courte si $d_2 = 0$; nous sommes dans le cas anti-persistant si $d_2 < 0$. Nous considérons le noyau polynomial défini par :

$$K(x) = \sum_{l=0}^{\tau} \alpha_l x^{2l} \text{ avec } |x| \leq 1 \quad (7)$$

et $K(x) = 0$ si $|x| > 1$. Ici $\tau \in \{0, 1, 2, \dots\}$ et les coefficients α_l sont spécifiés tels que $\int_{-1}^1 K(x) dx = 1$ (voir Beran et Feng (2002a et b) pour les détails de la méthode d'estimation).

3. Analyse Empirique de la série du Dow Jones

L'étude empirique porte sur la série logarithmique journalière du Dow Jones qui couvre une période historique allant du 26/05/1896 au 17/08/2006 ($n = 30292$). La mise en œuvre des divers tests nécessite que la série analysée soit stationnaire. Les résultats des tests de Dickey-Fuller (Dickey et Fuller (1981)), Phillips-Perron (Phillips et Perron (1988)), Schmidt-Phillips (Schmidt et Phillips (1992)), KPSS (Kwiatkowski, Phillips, Schmidt et Shin (1992)) et Elliott-Rothenberg-Stock (Elliott, Rothenberg and Stock (1996)) reportés dans le tableau 1 montrent que cette série est caractérisée par la présence d'une racine unitaire. La série est finalement différenciée pour obtenir les rentabilités (voir figure 1).

Figure 1 – Evolution de la série journalière du Dow Jones en logarithme et de la rentabilité

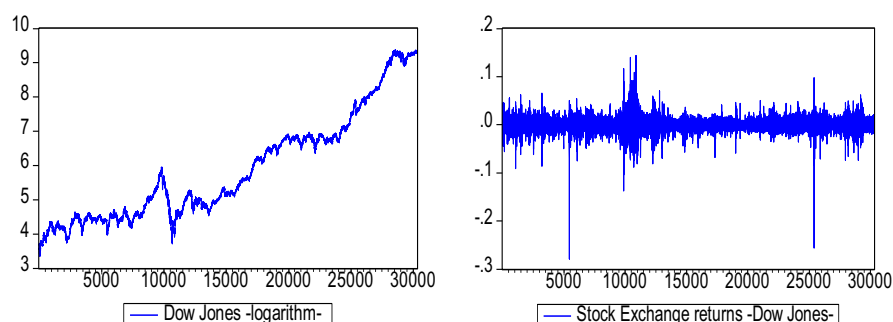


Tableau 1 – Tests de racine unitaire de Dickey-Fuller, Phillips-Perron et Schmidt-Phillips

Le test	Série en logarithmes		rentabilités	
Dickey-Fuller	(I)	-1.6316 (-3.4127)	(I)	-73.810 (-3.4127)
Phillips-Perron	(III)	-1.6323 (-3.4127)	(II)	-167.53 (-3.4127)
Schmidt-Phillips	$Z(Rho)$	-6.4786 (-18.1)	$Z(Rho)$	-27677.6873 (-18.1)
	$Z(Tau)$	-1.7970 (-3.02)	$Z(Tau)$	-159.0685 (-3.02)

Le test	La fenêtre	La méthode d'estimation spectrale							
		Noyau de Bartlett				Noyau spectral quadratique			
		Série en logarithmes		Rentabilités		Série en logarithmes		Rentabilités	
		(II)	(III)	(II)	(III)	(II)	(III)	(II)	(III)
KPSS	Newey-West	20.413 (0.463)	3.564 (0.146)	0.165 (0.463)	0.033 (0.146)	72.771 (0.463)	12.631 (0.146)	0.173 (0.463)	0.034 (0.146)
Série stationnaire	La fenêtre d'Andrews	0.7086 (0.463)	0.150 (0.146)	0.177 (0.463)	0.035 (0.146)	2.29 (0.463)	0.159 (0.146)	0.181 (0.463)	0.036 (0.146)
Elliott-Rothenberg-Stock	Newey-West	0.00272 (3.26)	0.0066 (5.62)	138.388 (3.26)	16.367 (5.62)	0.0028 (3.26)	0.0069 (5.62)	144.822 (3.26)	17.151 (5.62)
H_0 : racine unitaire	La fenêtre d'Andrews	0.0029 (3.26)	0.007 (5.62)	147.891 (3.26)	17.527 (5.62)	0.0029 (3.26)	0.0070 (5.62)	151.514 (3.26)	17.959 (5.62)

(I) : modèle sans constante ni tendance déterministe (à 5%)

(II) : modèle avec constance sans tendance déterministe (à 5%)

(III) : modèle avec constante et tendance déterministe (à 5%)

L'hypothèse de normalité des rentabilités est clairement rejetée (cf. tableau 2 et Figure 2). L'asymétrie constatée peut être le signe de la présence de non linéarités dans le processus d'évolution des rentabilités. Le diagramme de dispersion de la série (figure 3) ne se présente pas sous la forme d'un ellipsoïde régulier et confirme la non linéarité de la série. De plus, la série est hétéroscédastique d'après les résultats des tests de White et de Breusch-Pagan

reportés dans le tableau 3, puisque l'hypothèse nulle d'homoscédasticité est rejetée au seuil de 5%. Le résultat du test de l'hétéroscédasticité conditionnelle confirme ces résultats où la série des rentabilités de l'indice Dow Jones est caractérisée par la présence d'un effet *ARCH* très fréquemment rencontré dans les séries financières [$nR^2 = 1114.3201 > \chi_{0.05}^2(2)$].

Tableau 2 – Tests de normalité de la distribution des rentabilités

Skewness	Kurtosis	Stat. de J.B	Stat. de A.D	Stat. de K.S
-1.1208	38.2164	1571629	1015.4123	0.581 (0.092)

J.B est la statistique de Jarque-Bera,

K.S est la statistique de Kolmogorov-Smirnov,

A.D : la statistique de Anderson-Darling.

Les valeurs critiques de Kolmogorov-Smirnov sont données entre parenthèses sur la seconde ligne au seuil 5%

Figure 2 – Estimation de la densité par la méthode du noyau

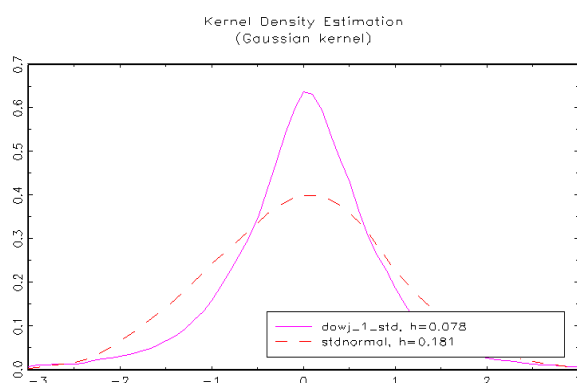
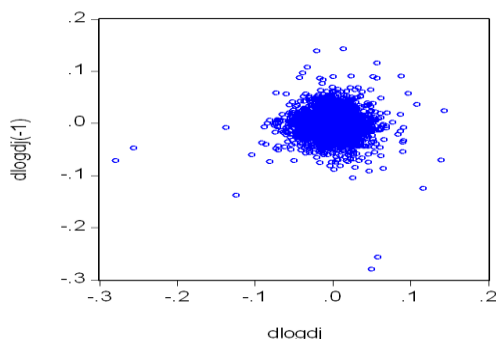


Tableau 3 – Tests d'homoscédasticité

Statistique de Breusch-Pagan	Statistique de White	Stat. de LM - ARCH(2) -
1481.7629	1370.9883	1114.3201

Figure 3 – Dispersion des variations successives des rentabilités boursières



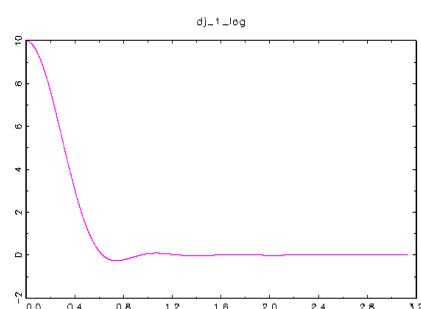
Au regard du tableau 4, l'hypothèse de marche aléatoire est clairement rejetée. En effet, les statistiques de Mizrahi (Mizrahi (1995)) et de *BDS* (Brock, Dechert et Scheinkman (1987)), qui testent la présence des dépendances de type linéaire ou non linéaire, sont strictement supérieures à la valeur critique au seuil de 5%.

Tableau 4 – Résultats des tests de BDS et de Mizrach sur la série des rentabilités

m	Les statistiques de BDS		Les statistiques de Mizrach
	Fraction of pairs	Standard Deviation	$T \leq 1000$
2	41.85412	37.41965	4.3802
3	53.06906	48.96358	3.1270
4	60.44449	57.82082	2.1684
5	67.01797	67.16815	2.1346
6	73.79956	78.60053	1.1964
7	80.97002	92.77881	1.7861
8	88.54294	110.1803	1.8503
9	97.17348	132.5872	1.9537
10	106.8643	161.0307	1.5520
11	117.9473	198.9349	-
12	131.0207	250.3263	-
13	146.1977	318.0141	-
14	163.8683	408.8996	-
15	184.3967	530.3570	-
16	208.5077	694.4048	-
17	237.1340	922.2850	-
18	271.0007	1242.474	-
19	311.1570	1689.988	-
20	359.0700	2325.839	-

La statistique de BDS est calculée par deux méthodes avec \mathcal{E} égale à 0.7.

Ces premiers tests font généralement ressortir la présence d'autocorrélations significatives différentes de zéro à court terme ; même si ces autocorrélations sont significatives d'un point de vue statistique, elles ne le sont pas d'un point de vue économique au sens où il est impossible d'exploiter ces autocorrélations pour établir des règles de spéculation conduisant à des profits anormaux. Les tests nous conduisent à rejeter l'hypothèse nulle d'absence d'autocorrélation, mais ne permettent nullement de déceler la présence d'une structure de dépendance à long terme. Devant cet état de fait, nous analysons le comportement cyclique des rentabilités en travaillant sur des horizons plus longs. En traçant le périodogramme de cette série (voir figure 4) (avec la fenêtre de Tuckey), nous notons que la densité spectrale est concentrée autour des faibles fréquences et tend vers l'infini lorsque la fréquence tend vers zéro. Il s'agit d'un signe de mémoire longue sur série stationnarisée qui reste à confirmer avec des tests statistiques.

Figure 4 – Le périodogramme des rentabilités du dow Jones

Pour cette raison, nous avons retenu diverses valeurs pour les ordonnées du périodogramme afin d'encadrer la racine carrée du nombre d'observations. Ce choix a pour objet d'examiner la stabilité des estimateurs lorsque le nombre d'ordonnées du périodogramme varie. D'après le tableau 5, il est évident que la série des rentabilités de l'indice Dow Jones est générée par un processus de mémoire longue. Quelques valeurs de la statistique de Student (avec une puissance de 0.8) sont strictement supérieures à la valeur critique au seuil de 5%. De plus, Le paramètre de mémoire estimé par la méthode d'Andrews et Guggenberger (2003) est positif et significatif. Le résultat d'estimation est très proche de ceux trouvés avec la méthode de GPH (Geweke et Porter-Hudak (1983)). Les rentabilités sont donc prévisibles à long terme : la présence d'une mémoire longue indique que les agents peuvent anticiper leurs rentabilités à un horizon suffisamment long. En effet, les mouvements observés apparaissent comme le résultat de chocs exogènes durables venant affecter le marché de New York, c'est-à-dire que le cours ne reviendra pas vers sa valeur fondamentale.

Tableau 5 – Résultats de l'estimation ARFIMA par les méthodes spectrales sur la série des rentabilités de l'indice Dow Jones

Fenêtre	Les ordonnées				
	$n^{0.4}$	$n^{0.5}$	$n^{0.6}$	$n^{0.7}$	$n^{0.8}$
GPH	-0.0489 (-0.5431)	0.0329 (0.6527)	0.0411 (1.4095)	0.0305 (1.7834)	0.0311 (3.0528)
Rectangular.	-0.0501 (-0.5045)	0.0114 (0.2061)	0.0322 (1.0023)	0.0282 (1.4961)	0.0293 (2.6160)
Bartlett	-1.1681 (-0.0670)	0.0132 (0.4104)	0.0304 (1.6382)	0.0253 (2.3253)	0.0271 (4.1915)
Daniell	-0.0694 (-0.9878)	0.0122 (0.3098)	0.0306 (1.3452)	0.0255 (1.9079)	0.0271 (3.4235)
Tukey	-0.0646 (-1.0326)	0.0115 (0.3288)	0.0309 (1.5249)	0.0260 (2.1866)	0.0277 (3.9173)
Parzen	-0.0614 (-1.1900)	0.5466 (0.0158)	0.0307 (1.8373)	0.0255 (2.6063)	0.0274 (4.7018)
B-priest	-0.0606 (-0.7878)	0.0103 (0.2406)	0.0317 (1.2713)	0.0270 (1.8500)	0.0285 (3.2790)
Andrews- Guggenberger	-	-	-	-	0.0294 (3.047)

Les valeurs entre parenthèses sont les statistiques de Student. En gras, figurent les valeurs statistiquement significatives.

Les résultats de l'estimation du modèle SEMIFARMA par la méthode du maximum de vraisemblance exact sont reportés dans le tableau 6. Après avoir estimé le trend déterministe, la fenêtre optimale et le critère de validation croisée par la méthode du noyau basée sur la méthodologie de Nadaraya-Watson (Nadaraya (1964) et Watson (1964)), les résultats indiquent que la série du Dow Jones étudiée est caractérisée par une mémoire longue : le paramètre d'intégration fractionnaire estimé est significativement différent de zéro. Ce résultat rejoint les conclusions trouvées avec les différents estimateurs des méthodes spectrales.

Tableau 6 – Estimation semi paramétrique par la méthode du maximum de vraisemblance exact

Maximum de vraisemblance exact	Estimation non paramétrique de la tendance déterministe	
	Critère CV optimal	\hat{h}_{opt}
$(2, 0.0423, 2)$ $t_d=12,5369$	0.9382	0.1384

CV : le critère de validation croisée et \hat{h}_{opt} : la fenêtre optimale

Nous remarquons que les résidus (figure 5) ne sont pas caractérisés par une distribution gaussienne (tableau 7) ; nous notons également le caractère leptokurtique des résidus (figure 6). L'asymétrie peut être le signe de la présence des non linéarités dans les résidus. Cependant, ces résidus peuvent être modélisés par les modèles *GARCH* car la présence d'un effet *ARCH* est confirmée par le résultat du test *ARCH LM* sur les résidus ($nR^2 = 20.1646 > \chi^2(1)$). Ainsi, le spectre des carrés des résidus (voir figure 7) est concentré autour des faibles fréquences et tend vers l'infini lorsque la fréquence tend vers zéro. Il est probable que la variance conditionnelle possède une structure de mémoire longue persistante. Les prix des actifs financiers présentent souvent des comportements de type hétéroscédastique avec de la persistance. C'est pour cette raison qu'on va examiner la variance conditionnelle de la série des rentabilités du Dow Jones, afin d'étudier la possibilité d'effet durable des chocs sur la volatilité.

Figure 5 – Evolution des résidus d'estimation

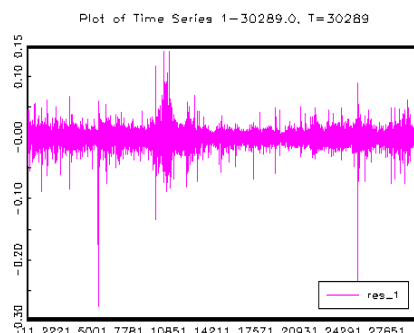


Tableau 7 – Caractéristiques essentielles de la série des résidus

Skewness	Kurtosis	Stat. de J.B	Stat. de <i>ARCH-LM</i>
-1.02	37.47	1505311.66	20.1646

Figure 6 – Estimateur à noyau de la densité de probabilité des résidus

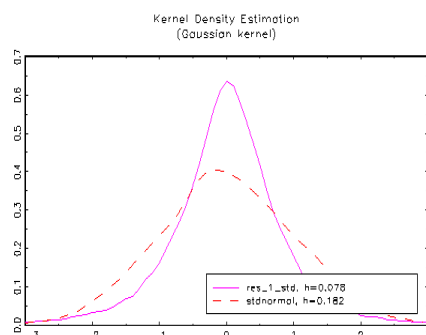
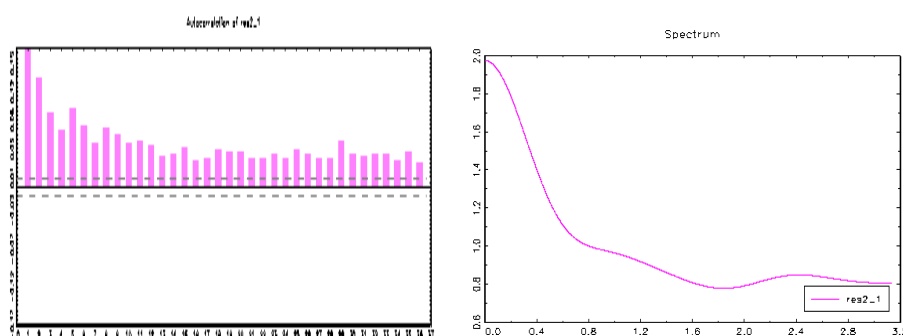


Figure 7 – Le corrélogramme simple et le périodogramme de la série des carrés des résidus



La détermination du processus suivi par notre série est ensuite estimée par la maximisation simultanée des processus de la moyenne et de la variance. Les tests nous permettent de juger de la solidité du modèle, d'une part, au niveau de l'absence de l'autocorrélation et de l'hétéroscédasticité, et d'autre part, de l'existence effective de mémoire longue et de la bonne modélisation de l'effet ARCH. Sur le plan pratique, on estime plusieurs formes de modèles pour différents retards : un SEMIFARMA (p, d, q) conjointement à un GARCH, FIGARCH, HYGARCH. Pour chacun de ces modèles, nous calculons les deux critères d'information d'Akaike (1970) et Schwarz (1978). L'estimation des différents modèles est faite seulement sur 30032 observations de l'échantillon total, pour pouvoir réaliser ultérieurement des comparaisons avec les prévisions des 260 observations restantes.

Tableau 8 – Estimation par la méthode du maximum de vraisemblance – l’algorithme de BHHH –

Les paramètres	SEMIFARMA-GARCH	SEMIFARMA-FIGARCH	SEMIFARMA-HYGARCH
$\hat{\phi}_1$	-0.1767 (-2.3134)	-0.2036 (-2.5513)	-
$\hat{\theta}_1$	0.2668 (3.7227)	0.5011 (0.7244)	0.1067 (12.3140)
\hat{d}_2	0.0326 (3.4389)	0.0231 (2.6932)	0.0226 (2.7058)
$\hat{\alpha}_0$	1.3564 (0.8644)	1.1025 (2.0876)	1.8451 (2.0039)
$\hat{\alpha}_1$	0.0895 (2.1494)	0.0612 (2.1511)	0.0567 (2.1765)
$\hat{\beta}_1$	0.0949 (4.1676)	0.1032 (3.9842)	0.1256 (2.4783)
\hat{d}	-	0.04311 (4.2745)	0.04420 (4.8655)
$\hat{\alpha}$	-	-	0.1592 (2.9252)
\hat{h}_{opt}	0.1384	0.1384	0.1384
<i>IMSE</i>	0.9382	0.9382	0.9382
<i>Akaike</i>	5.6842	5.6438	5.6395
<i>JB statistic</i>	3436.18	3407.22	3391.46
<i>Schwarz</i>	5.6903	5.6422	5.6416
<i>ARCH(1)</i>	2.8847	2.6512	2.4956

Les valeurs entre parenthèses sont les statistiques de Student

IMSE : Minimum Integrated Mean Square Error

Au regard du tableau 8, nous constatons que les critères d’information sont tous minimaux pour le modèle SEMIFARMA-HYGARCH et les coefficients de ce modèle sont hautement significatifs. De plus, les coefficients de mémoire longue relatifs aux équations de la moyenne et de la variance conditionnelles sont également significatifs et les résidus des modèles examinés sont caractérisés par l’absence de l’hétéroscédasticité conditionnelle où les statistiques de ARCH-LM sont strictement inférieures à la valeur critique de χ^2 au seuil de 5% à 2 degrés de liberté. Il est à noter que l’hypothèse de normalité des résidus des modèles testés est clairement rejetée car les statistiques de Jarque-Bera sont strictement supérieures à la valeur critique de la loi de χ^2 au seuil de 5% à deux degrés de liberté. Au regard du tableau 9, la série des résidus du modèle SEMIFARMA-HYGARCH ne présente aucune structure de dépendance où les statistiques de BDS sont strictement inférieures à la valeur critique 1.96.

**Tableau 9 – Résultats du test de BDS sur les résidus
du modèle SEMIFARMA-HYGARCH**

m	Les statistiques de BDS
2	0.2531
3	0.2844
4	0.3142
5	0.3578
6	0.4812
7	0.6740
8	0.8124

Afin de comparer les performances prévisionnelles des modèles proposés et celui de marche aléatoire, deux critères sont utilisés, l'erreur moyenne quadratique (QME) et l'erreur absolue moyenne (MAE) :

$$QME = H^{-1} \sum_{h=1}^H (\hat{Y}_{n-H+h} - Y_{n-H+h})^2 \quad (8)$$

$$MAE = H^{-1} \sum_{h=1}^H |\hat{Y}_{n-H+h} - Y_{n-H+h}| \quad (9)$$

où h est l'horizon de prévision et H est le nombre total de prévisions correspondant à l'horizon h sur la période prévisionnelle.

Le tableau 10 comporte les résultats de prévisions fournies par les différents modèles. Les critères QME et MAE donnent généralement les mêmes résultats. Nous remarquons que, quel que soit l'horizon de prévision, la modélisation avec la marche aléatoire est battue par tous les modèles testés. Nous constatons en général de bonnes capacités prévisionnelles du modèle SEMIFARMA-FIGARCH et tout particulièrement du modèle SEMIFARMA-HYGARCH à un horizon de 90, 180 et 260 jours. Nous avons calculé également des prévisions hors-échantillons (voir tableau 11), nous remarquons aussi que, quel que soit l'horizon de prévision, tous les modèles testés peuvent battre la marche aléatoire. Nous constatons de bonnes qualités prévisionnelles du modèle SEMIFARMA-HYGARCH à un horizon de 30, 90 et 180 jours. En effet, la marche aléatoire ne prend en compte que la mémoire de court terme de la série et néglige en conséquence totalement la mémoire de long terme. Nous nous apercevons que les qualités prévisionnelles des modèles à mémoire longue dans la variance conditionnelle donnent des prévisions de qualité supérieure à un horizon long. La prise en compte de cette mémoire longue implique que pour une prévision à long terme (180 -260 jours), nous utilisons non seulement la dernière valeur observée de la série mais également tout l'historique pondéré de cette série.

Afin de tester la qualité prédictive du modèle SEMIFARMA-HYGARCH avec le modèle de marche aléatoire, nous pouvons aussi utiliser plusieurs tests bases sur le test asymptotique, les tests des signes, de Wilcoxon, de Morgan-Granger-Newbold et de Meese Rogoff (tous les tests sont présentés par Diebold et Mariano (1995)). Les p-values fournies par le tableau 12 indiquent clairement que l'hypothèse nulle de l'égalité de la précision prévisionnelle de deux modèle est fortement rejetée car les p-values sont strictement inférieures à 0.05, c'est-à-dire, la qualité prévisionnelle du modèle SEMIFARMA-HYGARCH est supérieure à celle du modèle de marche aléatoire.

Étant donné que la série des rentabilités de Dow Jones se caractérise par une présence des dynamiques de long terme dans les équations de la moyenne et de la variance conditionnelle, et par un phénomène d'hétéroscédasticité, la modélisation SEMIFARMA-HYGARCH permet le calcul des meilleures prévisions à court terme et à long terme que le modèle de marche aléatoire. Il y a également une supériorité évidente sur tous les autres modèles pour des horizons longs.

Bien que l'écart entre les critères QME et MAE n'apparaisse pas très important, ce qui compte dans cette étude est le caractère systématique des modèles. En effet, les mouvements de cours apparaissent comme le résultat de chocs exogènes durables venant affecter le marché boursier américain ; en d'autres termes, les conséquences d'un choc seront durables, le cours ne reviendra pas vers sa valeur fondamentale précédente et le choc sera persistant à long terme. Nous pouvons alors penser, du fait de la prévisibilité à long terme des rentabilités, qu'il est possible *à priori* d'établir des stratégies rémunératrices sur le marché de New York.

Tableau 10 – Comparaison des qualités prédictives

Horizon	Critère	Marche Aléatoire	SEMIFARMA-GARCH	SEMIFARMA-FIGARCH	SEMIFARMA-HYGARCH
1 jour	QME	7.0328	6.0644	6.1632	6.1665
	MAE	3.2744	2.1432	2.0323	2.0342
2 jours	QME	7.1214	6.0648	6.0728	6.0813
	MAE	3.3328	2.1263	2.1344	2.1397
15 jours	QME	7.1488	6.6231	6.6420	6.6422
	MAE	3.3611	2.6302	2.6481	2.6486
30 jours	QME	7.2201	6.2003	6.2078	6.2076
	MAE	3.4027	2.8804	2.8902	2.8869
90 jours	QME	8.1746	7.0201	7.0194	7.0072
	MAE	3.4726	2.6327	2.6319	2.6289
180 jours	QME	8.3011	6.9303	6.9211	6.9072
	MAE	3.5814	2.4411	2.4207	2.3995
260 jours	QME	8.1233	6.7012	6.6981	6.6837
	MAE	3.4916	2.1354	2.1349	2.1219

Table 11 – Comparaisons des prévisions hors-échantillons

Horizon	Criteria	SEMIFARMA-GARCH	SEMIFARMA-FIGARCH	SEMIFARMA-HYGARCH	Random Walk
1 day	MSE	6.1232	6.1353	6.1422	7.0632
	MAE	2.1432	2.0323	2.0344	3.4811
2 days	MSE	6.1271	6.1452	6.1488	7.1568
	MAE	2.1824	2.1956	2.1397	3.6232
15 days	MSE	6.6401	6.6511	6.6823	7.2402
	MAE	2.6902	2.7105	2.7113	3.8600
30 days	MSE	6.2602	6.2588	6.2552	7.6314
	MAE	2.9460	2.9410	2.9400	3.5602
90 days	MSE	7.0911	7.0812	7.0253	8.2679
	MAE	2.7033	2.6963	2.6289	3.9601
180 days	MSE	6.9219	6.9008	6.8814	8.7246
	MAE	2.4903	2.4855	2.3836	3.9901

Tableau 12 – Comparaison des qualités prédictives : Test de Diebold-Mariano

<i>Forecast observations numbers</i>	S_1	S_2	S_3	MGN	MR
30	2.2856 (0.0223)	2.9212 (0.0035)	3.6303 (0.0003)	-7.2450 (0.0000)	-2.3343 (0.0211)

The P-Values are in parentheses, S_1 : Asymptotic test statistic, S_2 : Sign test statistic, S_3 : Wilcoxon test statistic, MGN: Morgan-granger-Newbold test statistic, MR: Meese-Rogoff test statistic. Truncation lag used for asymptotic test = 3, Truncation lag used for Meese-Rogoff test = 8.

4. Conclusion

Dans cet article, nous avons recherché la présence de mémoire longue dans la série des rentabilités de l'indice Dow Jones. Dans cette optique, nous avons proposé un modèle à mémoire longue semi-paramétrique appelé SEMIFARMA avec erreur GARCH hyperbolique. Nous avons mis en oeuvre la méthode du maximum de vraisemblance exact pour estimer cette classe de modèles en prenant en considération le phénomène de persistance de long terme au niveau de la variance conditionnelle. D'après les résultats, les chocs informationnels ont des conséquences durables sur la volatilité et le modèle SEMIFARMA-HYGARCH montre une supériorité évidente sur tous les autres modèles proposés pour des horizons longs. Plus précisément, les prévisions tirées du modèle à mémoire longue montrent une amélioration nette par rapport au modèle de marche aléatoire, et ce, à tous les horizons ; en conséquence, l'efficience faible des marchés financiers semble violée pour la série des rentabilités du Dow Jones étudiées sur longue période. Ainsi, les travaux récents portant sur la modélisation de la volatilité au moyen des processus FIGARCH ou HYGARCH semblent particulièrement prometteurs et peuvent apporter des éléments nouveaux à la compréhension du processus d'évolution des séries financières.

Références

- Akaike, H. (1970), Statistical Predictor Identification, *Annals of Institute of Statistical Mathematics*, 22, 203-217.
- Aloui C, Mabrouk S. (2010) Value-at-risk estimations of energy commodities via long-memory, asymmetry and fat-tailed GARCH models, *Energy Policy*, 38, 2326–2339
- Andrews, D. W. K., et P. Guggenberger (2003), A Bias-Reduced Log-Periodogram Regression Estimator for the Long-memory Parameter, *Econometrica*, 71, 675-712.
- Baillie, R.T., Bollerslev, T, and Mikkelsen, H.O. (1996). Fractionally integrated generalized autoregressive conditional heteroskedasticity. *J. Econometrics*, 74, 3-30.
- Barkoulas, J.T., Baum C.F. (1997), Long memory and forecasting in Euroyen deposit rates, *Financial Engineering and the Japanese Market*, 4, 189-201.
- Beran, J. (1999). SEMIFAR models—a semiparametric framework for modelling trends, long-range dependence and nonstationarity,” discussion paper No. 99/16, Center of Finance and Econometrics, University of Konstanz.
- Beran, J. and Feng, Y. (2002a). SEMIFAR models - A semiparametric framework for modelling trends, long-range dependence and nonstationarity. *Computat. Statist. Data Anal.*, 40, 393–419.
- Beran, J. and Feng, Y. (2002b). Iterative plug-in algorithms for SEMIFAR models -definition, convergence and asymptotic properties. *J. Computat. Graph. Statist.*, 11, 690--713.
- Beran, J., Ocker, D. (1999a) SEMIFAR forecasts, with applications to foreign exchange rates, *Journal of Statistical Planning and Inference*, 80, 137-153.
- Beran, J., Ocker, D. (1999b) Volatility of stock market indices - an analysis based on SEMIFAR models, Working paper, Department of Economics and Statistics, University of Konstanz, Germany.
- Bollerslev, T. (1986). Generalized Autoregressive Conditional Heteroskedasticity. *J. Econometrics*, 31, 307-327.
- Brock, W. A., Dechert, W. D., and Scheinkman, J. (1987). “A test for independence based on the correlation dimension.” Discussion Paper 8702, University of Wisconsin-Madison.
- Cardamone E., Folkinshteyn D. (2007). HYGARCH Approach to Estimating Interest Rate and Exchange Rate Sensitivity of a Large Sample of U.S. Banking Institutions, Working Paper, Temple University, Department of Finance, USA.
- Christodoulou-Volos, C.C Siokis, F. M., (2006), “Long Range Dependence in Stock Market Returns”, *Applied Financial Economics*, 16:1331-1338.
- Conrad C. (2010) Non-negativity conditions for the hyperbolic GARCH model, *Journal of Econometrics*, 157, 441-457.
- Davidson, J., (2004). Moments and memory properties of linear conditional heteroskedasticity models, and a new model. *Journal of Business and Economic Statistics*, 22, 16-29.
- Dickey, D. and W. Fuller (1979). “Distribution of the Estimators for Autoregressive Time Series with a Unit Root,” *Journal of the American Statistical Association*, 74, 427-431.
- Diebold, F.X and Mariano, R.S. (1995). Comparing Predictive accuracy. *Journal of Business and Economic Statistics*, 13, 3, 253-263.

- Engle, R.F. (1982). Autoregressive conditional heteroskedasticity with estimation of U.K. inflation. *Econometrica*, 50 987–1008.
- Fama, E. F., (1965), “Random Walks in Stock Market Prices”, *Financial Analysts Journal*, 21, 5, 55–59.
- Fama, E.F. (1970). Efficient Capital Markets: A Review of Theory and Empirical Work. *The Journal of Finance*, 25, 2, 383-417.
- Fama, E. F., (1998), “Market efficiency, long-term returns, and behavioral finance”, *Journal of Financial Economics*, 49, 283–306.
- Feng, Y., Beran, Y., and K. Yu (2007), Modelling financial time series with SEMIFAR–GARCH model, *IMA Journal of Management Mathematic*, 18, 395-412.
- Geweke, J., Porter-Hudak, S., (1983), The estimation and application of long-memory time series models, *Journal of Time Series Analysis*, 4, 221-238.
- Granger, C. W. J. and Joyeux, R. (1980). An introduction to long-memory time series models and fractional differencing. *J. Time Ser. Anal*, 1, 15-30.
- Grossman, S.J. (1976). On the Efficiency of Competitive Stock Markets Where Trades Have Diverse Information. *The Journal of Finance*, 31, 2, 573-585.
- Gursakal S. (2010), Detecting Long Memory in Bulls and Bears Markets: Evidence from Turkey, *Journal of Money, Investment and Banking*, 18, 95-104.
- Hall, P., Hart, J., 1990. Nonparametric regression with long-range dependence. *Stochastic Processes Appl.* 36, 339–351.
- Härdle, W., Mungo, J. (2008). Value at risk and expected short fall when there is long range dependence, Discussion Paper 2008-006. Humboldt-universität zu Berlin, Germany.
- Hosking, J. R. M. (1981), Fractional differencing. *Biometrika*, 68, 165–176.
- Hosking, J. R. M. (1996), Asymptotic distributions of the sample mean, autocovariances, and autocorrelations of long-memory time series. *J. Econometrics*, 73, 261–284.
- Jensen, M.C. (1978). Some anomalous evidence regarding market efficiency. *Journal of Financial Economics*, 6, 2-3, 95-101.
- Kasman, A., Kasman, S. , Torun E. (2009), Dual long memory property in returns and volatility: Evidence from the CEE countries' stock markets, *Emerging Markets Review*, 10, 122–139.
- Kwan, W., Li, W.K., Li, G. (2011), On the estimation and diagnostic checking of the ARFIMA-HYGARCH Model, *Computational Statistics and Data Analysis*, forthcoming.
- Kwiatkowski, D., Phillips, P., Schmidt, P., Shin, Y. (1992), Testing the Null Hypothesis of Stationary Against the Alternative of a Unit Root: How Sure are we that Economic Time Series have a Unit Root?. *Journal of Econometrics*, 54, 159-178.
- Lillo, F., Farmer, J.D. (2004) The Long Memory of the Efficient Market, *Studies in Nonlinear Dynamics & Econometrics*, 8, 3.
- McMillan, D.G., Kambouroudis D. (2009) Are RiskMetrics forecasts good enough? Evidence from 31 stock markets, *International Review of Financial Analysis*, 18, 117–124.
- Mizrach, B. (1995), A Simple Nonparametric Test for Independence. Working Paper 1995-23, Rutgers University, USA.

- Nadaraya, E.A. (1964), On estimating regression. *Theory probability and their applications*, 9, 134-137.
- Ñíguez, T.-M., Rubia, A. (2006) Forecasting the conditional covariance matrix of a portfolio under long-run temporal dependence. *Journal of Forecasting*, 25, 439-458.
- Phillips, P.C.B. and P. Perron (1988). "Testing for Unit Roots in Time Series Regression," *Biometrika*, 75, 335-346.
- Ray, B.K., Tsay, R.S., 1997. Bandwidth selection for kernel regression with long-range dependence. *Biometrika* 84, 791–802.
- Schmidt, P., & Phillips, PCB (1992). LM Test for a Unit Root in the Presence of Deterministic Trends. *Oxford Bulletin of Economics and Statistics*, 54, 257–287.
- Schwarz, G. (1978), Estimating the dimension of a Model. *Annals of Statistics*, 6,461-464.
- Tang, T.L., Shieh, S.J., (2006). Long memory in stock index future markets: a value-at- risk approach. *Physica A*, 437–448.
- Watson, G.S. (1964), Smooth regression analysis. *Sankhyā*, A26, 359-372.
- Wei, Y., Wang, Y., Huang, D. (2010) Forecasting crude oil market volatility: Further evidence using GARCH-class models, *Energy Economics*, 32, 1477–1484

Documents de Recherche parus en 2012¹

- DR n°2012 - 01 : Abdoul Salam DIALLO, Véronique MEURIOT, Michel TERRAZA
« Analyse d'une nouvelle émergence de l'instabilité des prix des matières premières agricoles »
- DR n°2012 - 02 : Emmanuel DUGUET, Christine Le CLAINCHE
« Chronic Illnesses and Injuries: An Evaluation of their Impact on Occupation and Revenues »
- DR n°2012 - 03 : Ngo Van LONG, Antoine SOUBEYRAN, Raphael SOUBEYRAN
« Knowledge Accumulation within an Organization »
- DR n°2012 - 04 : Véronique MEURIOT
« Une analyse comparative de la transmission des prix pour l'orientation des politiques publiques : le cas du riz au Sénégal et au Mali »
- DR n°2012 - 05 : Daniel SERRA
« Un aperçu historique de l'économie expérimentale : des origines aux évolutions récentes »
- DR n°2012 - 06 : Mohamed CHIKHI, Anne PEGUIN-FEISSOLLE, Michel TERRAZA
« Modélisation SEMIFARMA-HYGARCH de la persistance du cours du Dow Jones »

¹ La liste intégrale des Documents de Travail du LAMETA parus depuis 1997 est disponible sur le site internet : <http://www.lameta.univ-montp1.fr>

Contact :

Stéphane MUSSARD : mussard@lameta.univ-montp1.fr

